

Dan, By this evening

Brewster would like
you to review this.
If you'd like to make
any changes, I could
do it for you if you
let me know by 6 pm (I'll
be leaving at 6:30p).
The final draft needs to
be faxed to Sunil Paul
(fax sheet attached). Judy

WAIS Inc

Wide Area Information Servers
Aug 14, 1994

William Marriott
Sunil Paul
America Online
FAX 703-506-1942

Dear Bill and Sunil:

Here is a first cut at a functional specification for Lightning. This is not meant to be a full spec, rather a cut through all the commands to make sure we know how to respond to them. Through this excersize we changed the architecture somewhat to accomodate the needs of the system. Most notably, we generate the final HTML when the user wants to see an article or collection since these documents can change.

You will notice the name Will Tuthill on the spec. He will be doing a bunch of the coding on the project. Dan and I will still be actively involved in the technical as well as business sides of the agreement.

I would also like to actively pursue the final contract. I think all of us would rather just do the work, but we must get the contract in order. I believe that is between Sunil and myself. I can come up with a version if you like through my lawyers. (I hear yours are swamped). Until that contract says otherwise, all work and "property" developed for this project is owned by both companies without restriction.

On the personal front, my wife and I are going into the hospital tonight to have our first child! I will be taking a break of a week, and then in email and phone catchup mode for the next week after. Yippie! Therefore, I will not be in the phone call this week, but I hope this spec helps. Dan and Will will be available on the WAIS Inc side.

Sincerely yours,



Brewster Kahle

Lightning Functional Specification

Will Tuthill
Dan Aronson
Brewster Kahle
WAIS Inc
DRAFT 0.1
8/14/94

Lightning Functional Specification.....	3
Components	6
WAIS DBMS.....	6
HTML Template Format	7
Feed Handler	7
HTML Generator	8
Template DB	8
WAISgate.....	8
Feed Handler Implementation Overview	9
Installation.....	9
Data Structures	9
Returned Messages	10
State Variables	10
DBMS Commands From Feed Handler	10
Rainman Commands	11

This proposal is in response to: Thunder and Lightning Taking the Internet by Storm with AOL Publishing Tools by Sunil Paul and William Marriott, May 25, 1994, where the initial project was defined as:

Lightning -- A fast-track technology that lets any AOL IP publish to the Internet with little effort. The Lightning strategy is to position AOL's Rainman Plus as a platform from which to publish to Internet formats. Just as word processing applications gain market share by complying with their competitor's formats, Lightning will make Rainman Plus a more attractive interactive publishing platform. The target delivery of the product is September/October 1994. (See the attached screen shots of Mosaic and Rainman Plus for Time, Inc.)

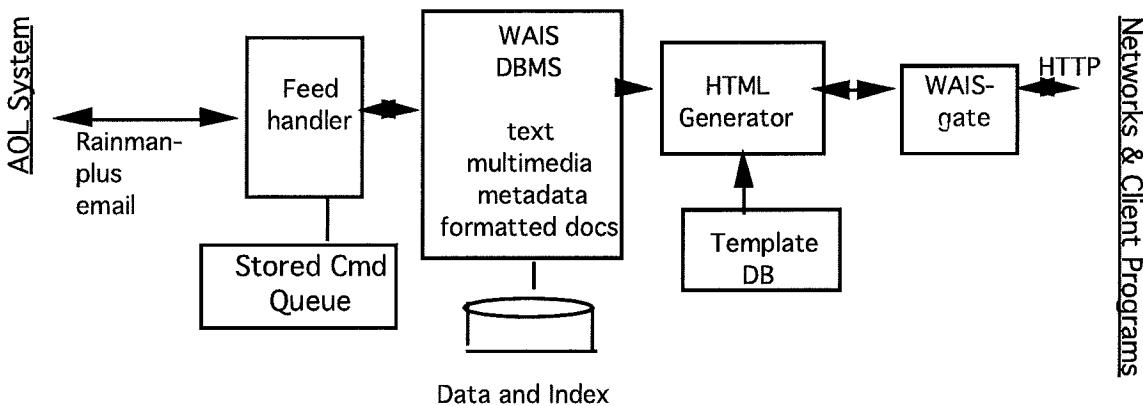
Features of the System:

- Integrated system for IP's to publish on Internet using Rainman system.
- Graphics and text integrated into a form will be displayed in a similar format HTML
- List boxes will be translated to a list of actionable hypertext buttons in HTML.
- Database searches will translated to form-based HTML entry boxes that can then be tied to WAIS searches.
- Information Providers will have the ability to select on a form by form level the ability to translate to HTML for Internet publication.
- There will be no provision for sales and credit-card transactions over the Internet with this translation.
- Text in a scroll box will be presented as a separate HTML page with the first x characters visible as a clickable hypertext button.
- Chat, Newsgroup/MsgBrd, software libraries, non English documents, and Surveys will not be supported.

This functional specification outlines the components and operation of the Lightning system. Please comment on anything related to this project. This draft concentrates on the different RAINMAN commands and is not a full spec by any means.

This module duplicates the AOL Information Provider services in an Internet service using HTTP and HTML (World Wide Web). This module does not handle discussion boards, message boards, conferencing or ties to other AOL services. In other words, this will build a similar service to the broadcast-oriented publishing services of AOL.

Since America Online personnel understand many of the implications of implementing Rainman-plus, we suggest these are our capabilities, but would need further work with the experts to find difficult areas.



Lightning Block Diagram

Components

The technical components of the Lightning system are code modules that run on a Sun Computer.

WAIS DBMS

This module stores the documents, indexes them, and serves them to users. This is built around the WAIS Server for Unix product by WAIS Inc.

Inputs: documents get added/replaced *Deleted*

Outputs: documents get read back to the feed hander and to the HTML generator,

Search results get passed to the HTML generator

State: Stores articles and collections as SGML documents (it may be more prudent to store it in some intermediate representation rather than SGML, but this document will talk about SGML).

Functions:

Search databases

Add documents

Retrieve documents

Index documents for searching

Delete documents

HTML Template Format

This is a format for expressing article and collection formats. This will be some sort of HTML but with place holders for pieces to be inserted from the documents.

Example: <to come>

Feed Handler

See the specification below.

Inputs: Rainman email messages

Output: email responses, SGML to WAIS DBMS, log file

State: Stored Command Queue

Installation: called from the aliases file of a Unix machine.

Functions:

Interpreting Rainman,

Reading SGML docs from WAIS DBMS into OBJECTs

Writing OBJECTs into WAIS DBMS in SGML

Add commands to Stored Command Queue (release) *initial response*

Modify commands in Stored Command Queue (context)

HTML Generator

This module takes an SGML document from the database and HTML Template to create a HTML document that can be passed to the user.

Inputs: SGML documents, HTML Templates (from the Template DB)

Output: HTML document

State: none

Template DB

This module stores HTML Template format documents.

Inputs: Hand created templates. Eventually, this could come from some tool.

Outputs: Templates go to the HTML Generator

State: set of template files.

Functions: All this "database" does is store the templates. Since they are numbered, this might just be a numeric filename in a particular directory. Therefore adding to it is easy and retrieving from it is fast and simple.

WAISgate

This module handles HTTP requests to the system and talks to the HTML generator. This is made up of an HTTPD implementation (assume NCSA's) and a set of PERL Scripts.

Inputs: HTTP from users

Outputs: HTML to users, requests to HTML generator

State: AOL help screens



Feed Handler Implementation Overview

All of the Rainman specific code will be in the email handling part of the feed handler. Therefore these additions to this module should be carefully designed and built with AOL involved. This section outlines the basic pieces of it.

Installation

This module is put as a command in the aliases file of a unix machine. Therefore a separate process gets created for each message that comes in. This has the benefit of being able to handle a large number of messages. It has the disadvantage that commands to the Stored Command Queue and the WAIS DBMS must be serialized.

Data Structures

The feed handler uses two basic datastructures for its operation: OBJECT, and RMID (Rainman Identifier). These objects are either created by the feedhandler or read from the database. Then they are operated on based on the Rainman commands and then stored back in the database.

OBJECT

type [ARTICLE COLLECTION]	<i>time</i>	<i>/* last modification time */</i>
author	/* the USER when object is created */	
backlinklist	/* id's of objects collections pointing to it */	
searchwords	/* list of extra words for the object */	
icon	/* its icon for display */	
title (mandatory field)	/* string */	
form	/* number of the form to use in displaying */	
date	/* the date associated with the object */	
datelineformat	/* where to display the date */	

ARTICLE

fieldvaluelist

database /* database for searching */

COLLECTION

maxlinks /* max number of links */

options /* ordering options */

~~newlink~~ /* link being inserted */

linklist /* ordered list of links */

Format?

RMID	/* makes up a unique name */
ip-id	/* info providers ID */
group-id	/* group ID */
[ARTICLE COLLECTION]	
ID	/* ID for article or collection */

Returned Messages

Email messages are returned to the originator with status, warnings, errors etc. Since this whole system is mirroring the AOL site, we will not be implementing all of the elaborate reporting mechanisms in AOL. On the other hand, it will be careful to not "crash and burn" on errors.

State Variables

These variables are reset when a new message comes in.

Warning_disposition

Context

Errors

Feedback

Group

Tryblock

User

Version

DBMS Commands From Feed Handler

These commands are used to manipulate the database.

```
int add(RMID name, OBJECT *object)  
int delete(RMID name)  
int read(RMID name, OBJECT **object)
```

all routines return error status

Rainman Commands

The list of Rainman commands are interpreted by the feed handler and translated into internal WAIS methods for operating on the data structures. This section describes what these different commands do.

ATTACHFILE

We ignore this command.

CLEARDATA

Clear the data from the specified field number or all field numbers in the data object.

COMMENT

Comments are ignored by the Feed Handler.

CONTEXT

Records in internal state which link number to process in the object's linklist. This can modify the object data structure, but it can also effect the Stored Command Queue if the schedule is changed. This extra feature would be helpful to put into the 2nd release of the software since it is the only place that manipulation of the Stored Command Queue is needed.

CREATE

Possibly ends previous create/modify (see END).

Initialize new object (creates RMID, OBJECT data structures.)

Creation is supported for objects of type ARTICLE and COLLECTION only.

DATA

Adds the given data to the specified field for the form in the field value list of the object.

Issue: What data formats do we need to support?

DATABASE

Sets the database attribute in the data object to the specified database or to no database when NONE is specified.

DATELINE

Adds the dateline and its format information to the dateline and dateline format attributes in the object.

DELIM

Changes the delimiter character used by the Rainman interpreter in the Feed Handler.

ECHO

Adds the text string specified in the ECHO command to the email buffer in the Feed Handler. The email buffer is mailed to the originator of the Rainman script when the Feed Handler is done processing the Rainman script.

END

Signals the conclusion of the creation or modification of an object. At this time the object stored internally is converted into SGML format and written into the WAIS database. Since END is optional, the functionality described here can also be triggered by the beginning of a new creation or modification.

If the "new link" slot is filled, then the link is inserted into the link list in the default ordering.

ENDTRY

Terminates the handling of a TRY command. If an error occurs between the TRY and ENDTRY command, then execution is resumed after the ENDTRY.

Issue: Is the ENDTRY an implicit END? We assume yes given the example script in the documentation.

Can we ignore TRY and ENDTRY until the next release?

ERRORS

Sets a flag internal to the Feed Handler to indicate whether or not all errors are considered fatal. Since the RAINMAN script has been processed at AOL already,

we do not expect to put in sophisticated error handling. Rather, if we get an error, we will send a message back to the originator with the rough status at the time of the error.

FEEDBACK

Sets the feedback level internal to the Feed Handler. The feedback level is used to determine what messages are added to the email buffer in the Feed Handler as processing of the Rainman script proceeds.

FIELDTEXT

Add the data to the specified field number in the representation of the object in the Feed Handler.

FORM

Selects the form from the Template Database and places a reference to this form in the representation of the object in Feed Handler.

FORMAT

Inserts this field into the object in the Feed Handler. This influences the HTML generator.

GROUP

Sets the Group variable which will be added to RMID's.

HEADLINE

Title
Sets the Headline slot in the current object.

HTEXT

This sets slots in the current object. Should it just set the Headline slot to this text? Should it be a special slot?

ICON

Sets the ICON slot in the current object.

ID

Sets the ID slot in the current RMID and Object. Calls the Create handler since this is just shorthand.

INSERT

Puts this link in the "new link" slot of the object. Since it is not clear where it should be positioned in the list. If there is a link already in the "new link" slot, then that old one is positioned into the list using the default ordering.

KILL

Schedules a delete of the links of an object, and the deletion of the backlinks in the objects that are pointed to. It also schedules the deletion of the object for one hour in the future.

LANGUAGE

Only English is handled.

MAP

This command is ignored.

MAXSIZE

This sets the maximum number of links in a collection slot in the current object.

MODIFY

This reads an object from the database into an object datastructure. Only ARTICLE and COLLECTION objects are handled.

OPTIONS

This sets options slots in the current object. *** more detail***

PASSWORD

This is ignored.

PLUSGROUP

This is ignored.

POSITION IN

This takes a link out of the "new-link" slot and puts it in the links list.

REFERENCE

*will add the appropriate links to
the collections*

This is ignored.

RELEASE

This adds an entry to the Stored Command Queue.

REMOVE

This adds an entry to the Stored Command Queue.

SEARCHWORDS

This adds these words to the searchwords slot in the current object.

SLUG

This is a confusing feature that adds an "alias" of sorts for an object. We would like to not implement this for the first release.

STOP

Not supported

TEXT

This adds text to the current object.

TEXTKEY

Not supported.

TITLE

This adds the text to the slot in the current object. Only TEXT is supported (not TYPED TEXT).

TRY

This sets a variable in the feed handler to not abort on fatal error. This variable is reset on ENDTRY.

USER

This sets a variable in the Feed Handler so that all objects have an author field filled in. Since all email messages will be coming from AOL, this is the only way to get user information for now.

VERSION

Only supports 2.0, otherwise it is a warning or some sort.

VIEWRULE

This is ignored for now until we understand it better.

WARNINGS

This sets a variable in the Feed handler on how to deal with warnings.